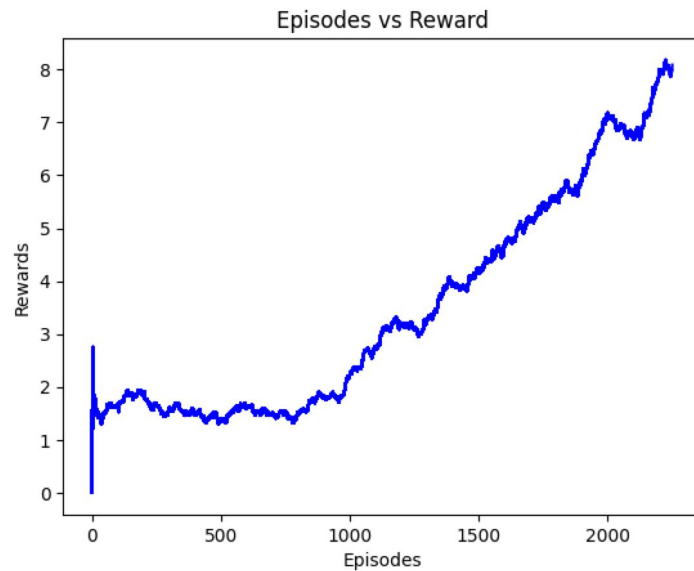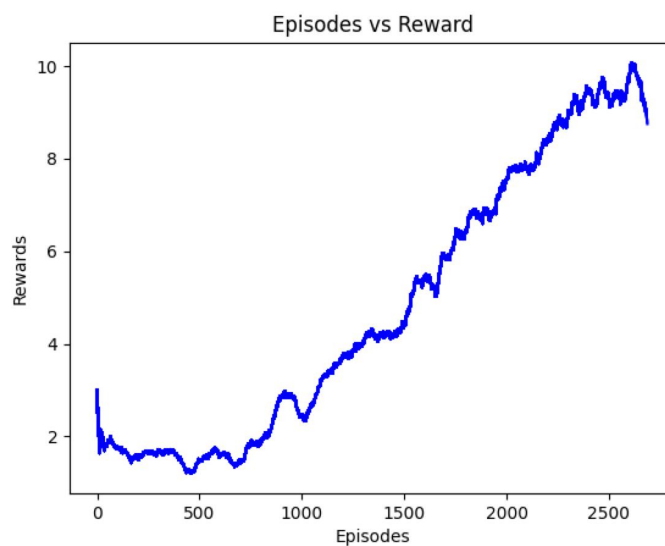**Name(s): Wenxuan Zhang**
Netid(s): wz47

Mean Reward Reached using DQN: 8.18

Plot of Mean Evaluation Reward vs. epochs for DQN model:

Mean Reward Reached using Double DQN: 10.07

Plot of Mean Evaluation Reward vs. epochs for Double DQN model:

Uploaded Saved DQN and Double DQN Models on Canvas: Yes

Uploaded your Agent.py and Agent_double.py files on Canvas : Yes

Provide a few sentences to analyze the training process and talk about implementation details:
The overtraining might occur after training to a certain episodes (most likely after 2500 episodes). During the training, the early stop strategy could be applied to avoid the overtraining issue. Pretraing could also be applied to increase the reward's value. For example, when the reward is greater than a certain threshold, we could save the model and train again with the saved one.
When implementing DQN agent and double DQN agent, one of the main difference is when computing the Q function in the next state. DQN uses policy net to get the next action, while double DQN uses target net to get the next action.

Extra Credit

1. **If you attempted the DQN LSTM Agent, give your implementation details. (Attempted to train the agent but failed.)**

2. **Videos: I have generated videos for DQN and double DQN, and uploaded them to canvas.**